Paper presented at the IEEE International Conference on Systems, Man & Cybernetics, 2889-2896, Washington D.C., Oct. 5-8, 2003.

Hidden Markov Models to Identify Pilot Instrument Scanning and Attention Patterns*

Miwa Hayashi
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA, U.S.A.
mhayashi@mit.edu

Abstract – *A previous study successfully demonstrated the potential usefulness of Hidden Markov Model (HMM) analysis techniques in analyzing pilots' eye-movement data to detect differences in scanning and attention patterns caused by display format changes. This paper focuses on assessing differences among pilots using the same display format. A flight simulator experiment was conducted with four pilots who had different levels of flight expertise. The analysis revealed variations in the HMM structures during the final descent segment: a 2-state HMM for the patterns of the least experienced pilot, a 3-state HMM for the two intermediate-level pilots, and a 4-state HMM for the most experienced pilot. The added "attitude-monitoring" state in the 4-state HMM reflected a flight technique well-known among experienced instrument pilots. HMM analysis methodology and interpretation issues are also discussed.*

Keywords: Aircraft display, attention, expertise level, eye movements, Hidden Markov Model (HMM), multiple tasks.

1 Introduction

When visibility is low and few or no visual cues from the outside are available, pilots must fly entirely by referring to cockpit instruments. In such low-visibility conditions, precise tracking of the assigned 3D path is critical, especially during near-ground maneuvers such as instrument approaches to ensure safe clearance over obstacles and terrain. Instrument flight can be thought of as concurrent multi-axis tracking on vertical (altitude), horizontal (course), and airspeed axes, where each axis has a different bandwidth and complex cross-coupling with the other axes [1]. Since pilots have finite cognitive resources, proper attention allocation among the tracking tasks is the key to successful overall tracking performance. There has been some controversy over how human operators attend to multiple tasks in general [2]. However, in a case such as instrument flight, where continuous and conscious tracking is required, the adequate working assumption is that pilots attend to the tasks in a serial manner rather than in a parallel manner. As Moray described in [2], one of the justifications is the nature of vision -- although there is evidence that pilots use peripheral vision, accurate instrument reading requires fixation. Because fixations occur one at a time, it is an adequate approximation that pilots attend tasks in a serial manner.

A number of researchers have analyzed pilots' eye-movement data to gain insight into pilots' basic attention management during instrument maneuvers. One study involving Hidden Markov Model (HMM) analysis was recently reported by the author's group [3]. The analysis revealed that a pilot's scanning and attention pattern changes when conventional digit-and-needle format gauges are replaced with digit-only format gauges for the airspeed indicator (ASI) and altimeter during simulated Instrument Landing System (ILS) approaches. HMM analysis demonstrated potential utility in interpreting pilots' eye-movement data.

However, there are still unsolved issues in using this analysis method. For instance, the analysis in [3] was focused on within-subject effects on one subject and did not include differences among pilots. Of course, different pilots are presumed to scan instruments in different ways and may require different HMM structures. This paper examines this fundamental issue in HMM analysis of pilot eye movements.

1.1 Differences Among Pilots

First, how do different individual pilots scan? Many investigators have reported large differences among pilots' scan patterns. Fitts, et al., for instance, noted marked differences in the eye-movement characteristics of 40 Air Force pilots [4]. For example, one pilot made 41 fixations

per minute on the artificial horizon during an instrument approach, while another pilot hardly used this instrument at all. They reported that there was also a trend that more experienced pilots made shorter fixations. Recently, Bellenkes, et al., measured eye movements of 12 novice pilots and 12 flight instructors flying various maneuvers on a simulator [1]. Their data revealed that the instructors tended to check the instruments more frequently. In addition, the instructors made more fixations on the flight variables that were not being maneuvered ("minding the store"), which may indicate a sign of good attentional flexibility. These findings suggest that, in addition to their own style difference, the pilots' experience level also may be a major factor in the different scanning strategies.

Prior studies, including those mentioned above, have been based on eye-movement statistics, such as fixation duration, look rate for each instrument, or transition probabilities between two instruments. These eye-movement statistics provide valuable information on how the pilots actually scanned the instruments. However, this type of analysis involves time averaging, and so it loses information regarding the sequence of instrument scans, which contains valuable information about pilots' attention-switching process. In addition, some instruments overlap in more than one tracking task, and that makes mapping back from fixation to each tracking task difficult. An example of an overlapping instrument is the attitude indicator (AI), which displays both pitch angle (for altitude and airspeed tracking) and bank angle (for course tracking). Investigators cannot determine which task the pilot was attending to when the pilot fixated on the AI.

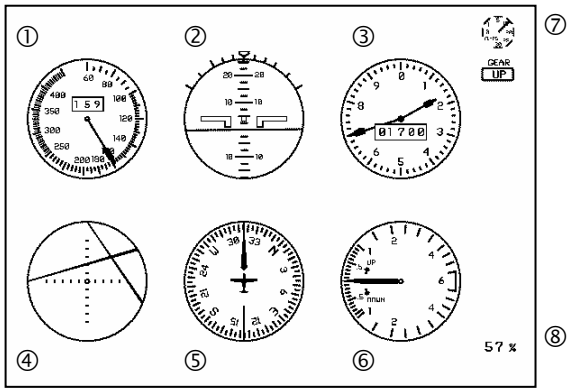
HMM analysis provides a solution for the ambiguity in tasks created by such overlapping instruments. It utilizes the sequential information of the pilots' instrument fixation to estimate a time-history of their attention switching among the different tracking tasks. The following section briefly reviews the concepts and procedures of HMM analysis.

1.2 Hidden Markov Model Analysis

Throughout instrument flight, pilots have to constantly crosscheck multiple instruments and interpret them because no single aircraft instrument completely describes the aircraft's behavior alone. HMM analysis exploits pilots' instrument crosscheck to solve the overlapping instrument problem. If an overlapping instrument is scanned along with other instruments that belong to a particular tracking task, then that instrument is more likely being used for that task.

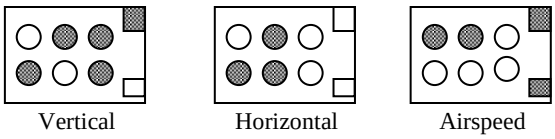
It is probably useful to consider a concrete example: Figure 1(a) shows a display used in the simulator experiment described in the next section. Figure 1(b) illustrates the instrument groups for each tracking task. Three instrument groups were identified and associated respectively with the vertical, horizontal, and airspeed-tracking tasks consistent with basic principles of attitude instrument flying (i.e., the instrument groups defined in [5] but with the ASI not in the vertical group and the AI in the airspeed group). We assume that during the flight, pilots switch their attention between these tasks, and fixate on the instruments associated with the current task. Note that Figure 1(b) indicates that the AI and the CDI are overlapping instruments because they appear in more than one task.

This concept can be translated into an HMM framework as shown in Figure 2. The HMM framework consists of two layers of stochastic processes. One is a hidden-state process, which is not directly observable but is assumed to follow first-order Markov process transition



①Airspeed Indicator (ASI), ②Attitude Indicator (AI), ③Altimeter, ④Course Deviation Indicator (CDI), ⑤Heading Indicator (HI), ⑥Vertical Speed Indicator (VSI), ⑦Flaps and Gear, and ⑧Thrust Indicator.

(a) Instruments



(b) Instrument grouping for each tracking task (indicated by the shaded instruments)

Figure 1. Basic-T display used in the simulator experiment

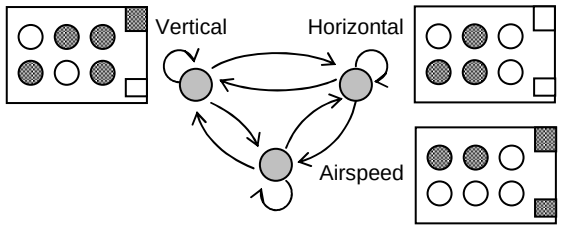


Figure 2. Three-state HMM

Table 1. HMM algorithms

A. The Forward Procedure (Scaled)

1. Initialization $\hat{\alpha}_1(i) = \pi_i\, b_i(o_1)$, $c_1 = \left[\sum_{i=1}^{N} \hat{\alpha}_1(i)\right]^{-1}$, $\hat{\alpha}_1(i) = c_1 \hat{\alpha}_1(i)$, $1 \le i \le N$

2. Induction $\hat{\alpha}_{t+1}(i) = \left[\sum_{j=1}^{N} \hat{\alpha}_t(j)\, a_{ji}\right] b_i(o_{t+1})$, $c_t = \left[\sum_{i=1}^{N} \hat{\alpha}_t(i)\right]^{-1}$, $\hat{\alpha}_t(i) = c_t\, \hat{\alpha}_t(i)$, $1 \le i \le N$, $1 \le t \le T-1$

3. Termination $\log\left[P(O \mid \lambda)\right] = -\sum_{t=1}^{T} \log c_t$

B. The Backward Procedure (Scaled)

1. Initialization $\beta_T(i) = 1$, $\hat{\beta}_T(i) = c_T$, $1 \le i \le N$

2. Induction $\beta_t(i) = \sum_{j=1}^{N} a_{ij}\, b_j(o_{t+1})\, \hat{\beta}_{t+1}(j)$, $\hat{\beta}_t(i) = c_t \beta_t(i)$, $1 \le i \le N$, $t = T-1, \ldots, 1$

C. The Baum-Welch Method

$$\xi_t(i,j) = \frac{\hat{\alpha}_t(i)\, a_{ij}\, b_j(o_{t+1})\, \hat{\beta}_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \hat{\alpha}_t(i)\, a_{ij}\, b_j(o_{t+1})\, \hat{\beta}_{t+1}(j)}, \qquad \gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j),$$

$$\bar{\pi}_i = \gamma_1(i), \quad \bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad \bar{b}_j(k) = \frac{\sum_{t=1,\, s.t.,\, o_t = v_k}^{T-1} \gamma_t(i)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad 1 \le i, j \le N$$

D. The Viterbi Algorithm (Logarithm Form)

1. Preprocessing $\tilde{\pi}_i = \log(\pi_i)$, $\tilde{a}_{ij} = \log(a_{ij})$, $\tilde{b}_i(o_t) = \log[b_i(o_t)]$, $1 \le i, j \le N$, $1 \le t \le T$

2. Initialization $\tilde{\delta}_1(i) = \tilde{\pi}_i + \tilde{b}_i(o_1)$, $\psi_1(i) = 0$, $1 \le i \le N$

3. Recursion $\tilde{\delta}_{t+1}(j) = \max_{1 \le i \le N} [\tilde{\delta}_t(i) + \tilde{a}_{ij}] + \tilde{b}_j(o_{t+1})$, $\psi_{t+1}(j) = \arg \max_{1 \le i \le N} [\tilde{\delta}_t(i) + \tilde{a}_{ij}]$, $1 \le t \le T-1$, $1 \le j \le N$

4. Termination $q_T^* = \arg \max_{1 \le i \le N} [\tilde{\delta}_T(i)]$

5. Backtracking $q_t^* = \psi_{t+1}(q_{t+1}^*)$, $t = T-1, T-2, \ldots, 1$

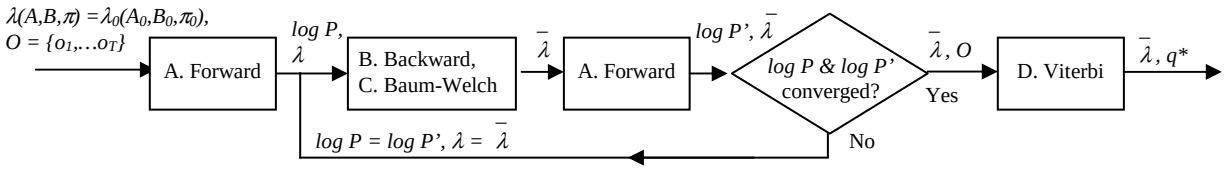


Figure 3. HMM computation flow

rules. The other is an observation-symbol process, which is physically observable and has a certain probability distribution depending on the current hidden state. In our case, the task switching (arrows in Figure 2) was modeled as the hidden-state process, and the instrument fixation corresponded to the observation process. Note that many researchers have calculated the transition probabilities among instrument fixations, which is equivalent to computing the Markov matrix of the fixations. In this analysis, however, the pilot's task was assumed to be the Markov process, not the pilot's fixations.

A model structure of an HMM is defined by a set of three parameter matrices. The initial state probability distribution matrix, $\pi = \{\pi_i\}$ $(1 \le i \le N)$, gives the probability that the hidden state, i, is the initial state; the state-transition probability distribution matrix, $A = \{a_{ij}\}$ $(1 \le i,j \le N)$, gives the transition probability from the hidden state, i, to the other state, j; and the observation symbol probability distribution matrix, $B = \{b_j(k)\}$ $(1 \le j \le N,\ 1 \le k \le M)$, gives the probability of the observation symbol, v_k, within the given hidden state, j. Detailed derivations of the HMM parameter estimation algorithms are described in [6]. Table 1 and Figure 3 summarize these algorithms and the computation flow.

The general computation procedure is as follows (Figure 3): it starts with an observation-symbol sequence (i.e., the fixation sequence data), $O = \{o_1, \ldots, o_T\}$, and the initial conditions for the HMM parameters, $\lambda_0 = (A_0, B_0, \pi_0)$. These initial conditions are determined heuristically. The estimation results are usually more sensitive to the choice of B_0 than that of A_0 or π_0. Especially the balances among b_j*(k=overlapping instruments)* seem to determine the bias on the Viterbi algorithm outcomes, and therefore should be chosen carefully. For example, if $b_{vertical}(AI) >> b_{horizontal}(AI)$, the Viterbi algorithm tends to pick the vertical task for AI fixations. Note that these parameters, including the initial conditions, are subject to the constraints: $\sum_j[a_{ij}] = 1$, $\sum_k[b_j(k)] = 1$, and $\sum_i[\pi_i] = 1$. When the backward procedure and the Baum-Welch algorithm are completed, $logP(O|\lambda)$ is computed by the forward procedure with the new λ, and checked for convergence. It usually does not take more than 3 to 4 loops for $logP(O|\lambda)$ to converge to within 1%. Finally, the Viterbi algorithm computes the estimated hidden-state sequence, $q^* = \{q_1, \ldots, q_T\}$. These estimates can be compared with the pilot's verbal reports, using a process that is described below.

Unlike other pattern recognition applications such as speech recognition, there is no training data available for the case of instrument flying analysis. Instead, the pilot's verbal reports of the attended tasks are used to train the

HMM. During the simulation flights, each pilot is asked to verbalize instrument readings (e.g., "1500 ft," "left of the course") or current intentions (e.g., "too low, climb a little," "turn to the right") as much as possible. These reports indicate the tracking tasks attended to at each moment, and are later converted to the corresponding tasks to be used as the training data. These training data are, in a way, run together with the actual observation data (the eye-movement data) in this HMM analysis, and when the training is complete, the estimation process is also complete. If an HMM estimated task matches the verbal report at the reported time or within ± 1 second, the report is considered to "match" the estimate. The verbal reports regarding single-parameter instruments (e.g., ASI, altimeter, HI) usually match, and thus are omitted from the analysis. Only the reports related to the overlapping instruments (i.e., the AI and CDI) are included in this verification process. If the rate of match is not satisfactory, the initial conditions are modified accordingly, especially those of $b_j(k=\textit{overlapping instruments})$, and the process is repeated until there is no further improvement.

The next section describes a flight simulator experiment conducted to examine the differences in the HMM structures among the four pilots who participated.

2 Method

2.1 Simulator Experiment

A fixed-base flight simulator configured with Boeing 757-200 flight dynamics was used. In the simulator, an aircraft control column was provided in front of the pilot, and a throttle lever at the right side of the pilot. The flaps, gear, and trim switches were provided on the yoke of the control column. The Basic-T display shown in Figure 1(a) was shown on a 17-inch computer monitor in front of the pilot. The display had a subtended visual angle of 23.0° horizontally and 17.6° vertically at a viewing distance of 30 inches. The six large instruments were 5.2° in diameter, with 1.4° horizontal and 1.7° vertical separations.

Four subjects of different expertise levels participated in the simulation. The pilots were numbered by order of experience. Pilot 1 was an instrument-rated Private Pilot with 250 hours of total and 60 hours of instrument time. Pilot 2 was a Certified Flight Instructor-Instrument with 700 hours of total and 100 hours of instrument time. Pilot 3 was a former military anti-submarine patrol aircraft pilot with 1050 hours of total and 225 hours of instrument time. Pilot 4 was a former airline pilot with 3500 hours of total and 350 hours of instrument time. Pilot 4 had flown in various air transporter aircraft including the Boeing 757, whose flight dynamics were simulated in this experiment.

The scenario simulated was an ILS approach. No out-the-window view was presented; i.e., all flights were made

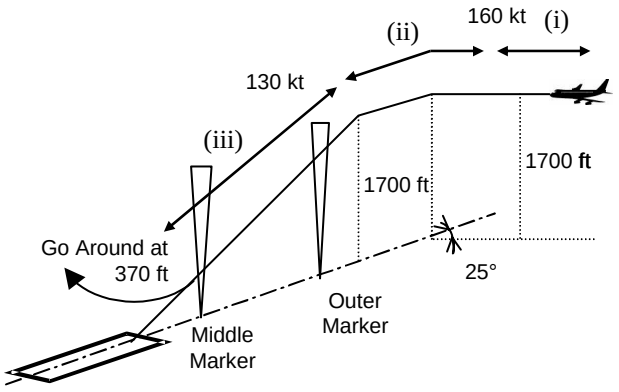


Figure 4. ILS approach scenario

referenced only to the instruments. Figure 4 shows a schematic of this approach. The aircraft was initially positioned on the left side of the localizer course. This created three flight segments each with somewhat different task requirements: (i) straight and level, (ii) intercept, and (iii) final descent. In segment (i), pilots were instructed to maintain 1700 ft and 160 knots. When the localizer needle moved in, the pilots started a left turn and this started segment (ii). This segment involved configuration changes, such as lowering gear and flaps, and slowing to 130 knots. When the localizer was centered and the glide slope needle started to move in, the pilots began the descent, and this started segment (iii). The pilots were instructed to keep the localizer and glide slope needles centered and maintain 130 knots. When the aircraft came down to 370 ft, a go-around was initiated and data collection ended for that approach.

During each approach, the pilots' eye-movement data were collected with a head-mounted eye-camera (RK-726PCI/RK-620PC, ISCAN Inc., Burlington, MA) and a magnetic head tracker (InsideTRAK, Polhemus, Colchester, VT) at the rate of 60 Hz. Flight variables were also recorded at 1 Hz. In addition, the pilots' verbal reports of the current intention or instrument reading were recorded on a videotape for later HMM analysis. Each pilot flew three data-collection approaches. Before the data-collection approaches, each pilot received a briefing and made several practice approaches, at least one of which was made while wearing the eye-tracking headgear and making verbal reports for familiarization.

2.2 Data Analysis

From the eye-movement data, intersection coordinates of the pilots' line-of-sight and the display plane were computed. Data associated with saccades were eliminated by omitting the data points that failed to meet the fixation criteria (i.e., staying within an ellipse of 5 pixels horizontally, 3 pixels vertically, for 2 consecutive samples). Since the line-of-sight intersections for each instrument were well separated, they could be easily clustered for the associated instrument by simple algorithms such as a K-

means clustering segmentation [7]. Intersections staying within the same instrument for less than 0.5 seconds were considered a single fixation. Likewise, intersections remaining for more than 0.5 seconds were counted as two fixations; those staying for more than one second were three fixations; and so on.

HMM analysis was applied separately to each pilot's data. In addition, since each flight segment had different task requirements, each segment was analyzed separately. (Note that this procedure is different from that in [3], where eye-movement data from all three segments of each approach were processed together.) The simplest way to estimate the parameters that maximize $P(O_1,O_2,O_3|\lambda)$, where O_r denotes the fixation sequence data from that segment of the r-th approach $(1 \le r \le 3)$, was to combine O_1, O_2, and O_3 to make one long data string. This process connected two data points that were not originally connected, but the effects were considered minimal because these were only two points in the long data string, and the model structure we used did not have any prohibited hidden-state transitions.

3 Results

First, the 3-state HMM in Figure 2 was applied to the analyses of all segments of all pilots. Table 2 shows the resulting rates of match with the verbal reports. Comparison showed high match rates of more than 80% in all cases, except segment (iii) for Pilot 4. The missed detections in this segment occurred mostly at the points when the pilot looked at the AI for a relatively long time (2-7 seconds) with occasional looks at the CDI, and verbally reported a pitch-related task (vertical task). The pilot also frequently looked at the heading indicator (HI), and the combination of viewing those three instruments tended to make the HMM estimate these portions as a horizontal task, which conflicted with the verbal reports. The interview with the pilot later revealed that, at these points, the pilot checked both pitch and bank to maintain stable path tracking on the final descent. Based on this information, a new hidden state, namely an "attitude-monitoring" task associated with long fixation on AI and occasional CDI, was added to make an alternative 4-state HMM (Figure 5). The re-computation with this model improved the match rate to 31/35 (89%).

Table 2. Rates of matches between verbal reports and the HMM hidden-state estimates (number of matches / number of total verbal reports) when the 3-state HMM was applied.

	Segment (i) Straight Level	Segment (ii) Intercept	Segment (iii) Final Descent
Pilot 1	13/13 (100%)	23/25 (92.0%)	55/63 (87.3%)
Pilot 2	27/27 (100%)	39/43 (90.7%)	37/41 (90.2%)
Pilot 3	28/28 (100%)	26/32 (81.3%)	41/45 (91.1%)
Pilot 4	13/15 (86.7%)	29/31 (93.6%)	22/35 (62.9%)

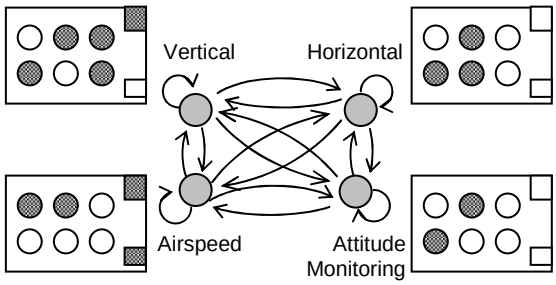


Figure 5. Four-state HMM

This fourth state may appear similar to the vertical task but is different due to the much longer fixation on the AI. This state means that the pilot was not actively controlling the aircraft. The aircraft was presumably already stabilized, and the pilot had found the "target" pitch and bank angles to maintain this stable tracking. In this case, crosschecking with the peripheral instruments may not be necessary at least for a while. This is a well-known technique among experienced instrument pilots. This fourth state did not appear in the other segments for Pilot 4.

Segment (iii) may have lead to interesting differences of attention strategy among pilots. In Table 3, the time percentages for each task estimated by the 3-state HMM, or an alternative HMM if any, within segment (iii) for each pilot are listed. Pilot 4's relatively low RMS flight technical errors and standard deviations of pitch and bank angles support the inference made above for this pilot. The table also indicates that Pilot 4 spent only 5% of the time on the airspeed task during this final descent segment, probably because the airspeed was stable enough and did not require much control. Therefore, an alternative model would be a 3-state HMM without an airspeed task, depending on the purpose of the analysis. Here, the fact that the pilot still kept checking the airspeed occasionally is important, and thus the airspeed tracking state is left in the model.

Table 3 also shows that Pilot 1 spent only 2.4% of the segment (iii) flight time on the airspeed task. It was, however, a different story from Pilot 4's case. The larger pitch and bank standard deviations and larger RMS flight technical errors indicate that the flight required more effort. The RMS airspeed error was relatively large, but the ASI was looked at for only 1.5% of segment (iii). The throttle was used several times, but almost always without an accompanying fixation on the ASI. It implies that the pilot was using the throttle mainly for lift control rather than for the airspeed control. This suggests merging the airspeed task into the vertical task to make a 2-state HMM (Figure 6). The estimated task percentages with the 2-state model during segment (iii) are shown in Table 3.

Similar attention characteristics of Pilot 1 also appeared in segment (ii). The airspeed tasks accounted for 6% of the total time in segment (ii), and the ASI was fixated

Table 3. Estimated task time percentages and other data in segment (iii) for all pilots

	Task time percentages (3-state HMM)			Task time percentages (Alternative HMM)				RMS flight technical error			Control surface statistics		Instrument fixation percentages					
	Vert. [%]	Horz. [%]	Aspd. [%]	Vert. [%]	Horz. [%]	Aspd. [%]	Monit. [%]	From glide slope [deg]	From localizer [deg]	From 130 knots [knts]	Pitch mean (std.) [deg]	Bank mean (std.) [deg]	ASI ① [%]	AI ② [%]	Alt ③ [%]	CDI ④ [%]	HI ⑤ [%]	VSI ⑥ [%]
Pilot 1	54.2	34.8	2.4	62.1	29.8	---	---	0.47	0.82	14.2	-5.33 (4.58)	-0.93 (5.61)	1.5	10.7	6.0	22.4	18.5	10.8
Pilot 2	47.5	22.9	12.9	---	---	---	---	0.38	0.28	15.3	-4.17 (1.70)	0.55 (2.86)	9.0	25.3	12.4	8.3	7.8	5.4
Pilot 3	51.6	21.4	12.7	---	---	---	---	0.30	0.53	6.6	-4.29 (1.95)	0.09 (2.94)	7.4	25.2	5.7	8.8	9.0	2.2
Pilot 4	9.3	80.0	5.0	9.1	43.4	5.0	31.9	0.11	0.13	7.1	-3.76 (0.63)	0.04 (1.74)	3.1	46.9	1.2	6.6	12.0	1.9

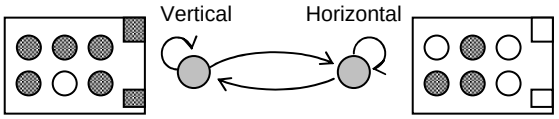


Figure 6. Two-state HMM

less than 1% of the time. The same 2-state HMM was applied for this segment, and the resulting time percentages are shown in Table 4. The 2-state model in segments (ii) and (iii) for this pilot could be a result of high workload demand during the localizer intercept and the final descent. In segment (i), however, where the workload is assumed lower than the other segments, the ASI was more often fixated, for 3% of the segment (i) time. As in segments (ii) and (iii), the pilot used the throttle for lift control rather than for airspeed control. Thus, a slightly modified 3-state model was applied to segment (i), and the results are shown in Table 4. The airspeed task was kept in the model this time because it conveys the important fact that, in this segment, the pilot occasionally looked at the ASI.

Table 4. Estimated task percentages and ASI fixation time percentages in segments (i) and (ii) for Pilot 1

	Task time percentages (3-state HMM)			Task time percentages (Alternative HMM)			Fixation percent.
	Vert. [%]	Horz. [%]	Aspd. [%]	Vert. [%]	Horz. [%]	Aspd. [%]	ASI ① [%]
Seg. (i)	62.8	28.5	6.2	68.5*	17.4	2.9 *	2.9
Seg. (ii)	17.2	54.2	5.5	23.9	54.1	---	0.2

* : Thrust Indicator was included in the vertical task, and not in the airspeed task.

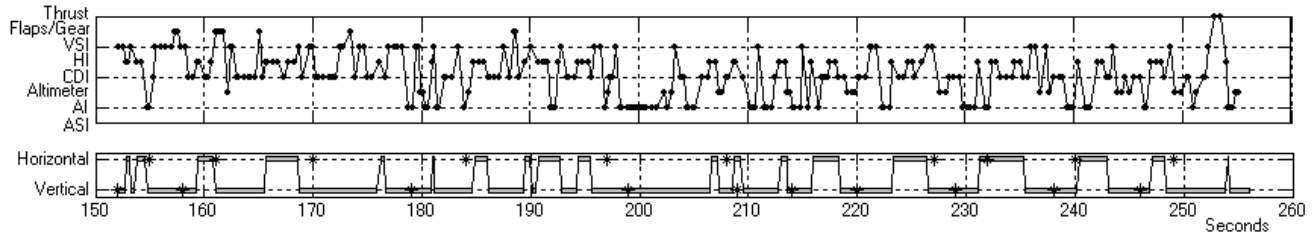
The data for Pilot 2 and Pilot 3 seemed to fit the original 3-state model for all of the segments. At least, there were few conflicts between the verbal report and the 3-state model estimation results. The airspeed tracking task accompanied by fixations on the thrust indicator occurred reasonably often. The fourth state that appeared in segment (iii) for Pilot 4, that is, the long fixations on the AI with occasional looks at the CDI, was not observed in any segment for these pilots. The 3-state model described the data of these pilots well, and therefore no modification of the model was attempted.

Figure 7 shows the instrument fixation data and associated HMM estimation results in segment (iii) of the last approaches each subject made.

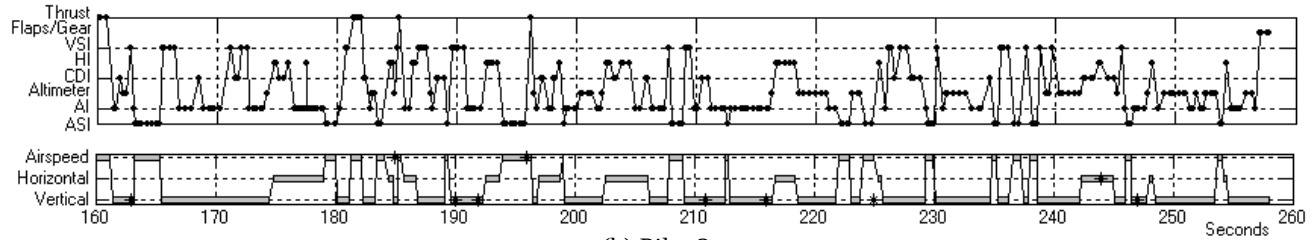
Finally, the look rates (fixations per second) and mean fixation durations (seconds) for all instruments were examined to compare with prior studies. In segment (ii), where maneuvering the aircraft was required to intercept the localizer, the look rates and mean fixation durations of the pilots ranged from 1.23 to 1.71 fixations/sec., and from 0.40 to 0.54 sec., respectively, and both variables showed correlations with the pilots' instrument time (0.0016 fixations/sec. increment per each instrument hour, $p = 0.039$, $R^2 = 0.9236$, and 0.0005 sec. decrement per each instrument hour, $p = 0.042$, $R^2 = 0.9178$, respectively). The more experienced pilot fixated on instruments more often and for shorter duration, which agrees with the prior findings [1], [4]. In segments (i) and (iii), however, no correlation with the pilots' expertise level was found.

4 Discussion

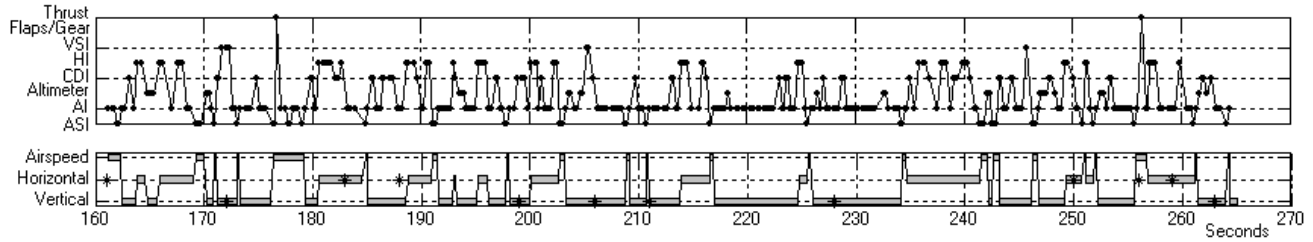
The HMM analysis was started with the 3-state HMM structure shown in Figure 2, and some variations of the model were derived that described the data of each pilot during the course of the analysis. These model derivations were based on simple observations. It would be convenient if there were an automatic and objective way to detect possible model structures for the given data. However, even if an automatic detection program were used, it would still be important for the researcher to be able to provide a



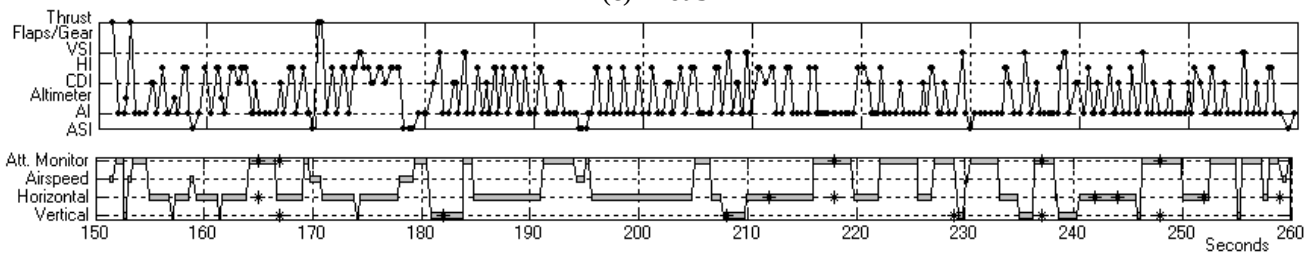
(a) Pilot 1



(b) Pilot 2



(c) Pilot 3



(d) Pilot 4

Figure 7. (a)-(d): The fixation sequence (top) and estimated task sequence (bottom) from segment (iii) in the last approach each pilot made. The asterisk marks on the bottom plots indicate the tasks implied by the pilot's verbal reports. In (d), where two asterisks appear at the same time point, the lower asterisks indicate the task originally implied for the 3-state HMM estimation; these were re-interpreted as the attitude-monitoring task when the 4-state HMM was applied.

reasonable explanation for each addition or removal of the hidden states. For instance, it is always possible to divide one of the hidden states into multiple substates and still fit the data with it. For example, the fourth state, the attitude-monitoring task, for Pilot 4 could be considered a substate of the vertical tracking task. However, researchers should consider whether it is really necessary to do so, as increasing the number of hidden states will incur a significant cost in computational efficiency. The amount of data required for adequate estimation quality will also increase as the number of hidden states increases. If multiple models describe the given data equally well, it is better to stay with the model that has the smallest number of hidden states.

One of the criteria used in this paper to justify adding states was the match rate of the verbal reports. The fourth state was added to segment (iii) for Pilot 4 because it improved the match rate significantly. The other criterion used was to preserve certain characteristics of the data. For instance, the airspeed tasks in segment (i) for Pilot 1 and segment (iii) for Pilot 4 were kept in the model because it was important to know that the pilots still looked at the ASI occasionally. If these characteristics were not of interest, then the airspeed task could have been removed. Use of this criterion is mostly up to the researcher's discretion.

In this analysis, the numbers of hidden states and the pilots' expertise level showed an interesting positive correlation, and that was not a coincidence. The 2-state

model indicated that one of the tracking tasks was dropped because of the high workload that the less-experienced pilot had, while the added attitude-monitoring state in the 4-state model indicated extra time the experienced pilot had due to the well-stabilized aircraft behavior during the final descent segment. The result implies an interesting circular effect that skilled pilots fly well not only because they track and stabilize well, but also because the extra time this creates gives them more attentional resources so that they can more quickly respond. This implication would not have been revealed by analysis of eye-movement statistics alone.

5 Conclusion

A number of previous studies have reported large differences among pilots in how they scan instruments. In this paper, eye-movement data from four pilots with different expertise levels flying simulated ILS approaches were collected, and analyzed for their attention patterns using HMM analysis. The data from two intermediate-level pilots were described well by the 3-state HMM, including vertical-, horizontal-, and airspeed-tracking tasks, for all segments. The localizer intercept and the final descent segments data of the least experienced pilot indicated that the pilot dropped the airspeed-tracking task, and thus the pilot's data fit well with the 2-state HMM, including vertical- and horizontal-tracking tasks. The data from the final descent segment of the most experienced pilot were best described by the 4-state HMM, which includes vertical-tracking, horizontal-tracking, airspeed-tracking, and attitude-monitoring tasks. The last task is consistent with a flight technique well-known among experienced instrument pilots. The results showed a positive correlation between the number of hidden states and the pilots' expertise level.

Acknowledgement

This study was supported by FAA Office of the Chief Scientific and Technical Advisor for Human Factors, AAR-100, working in cooperation with the FAA Transport Airplane Directorate, AMN-111. The author thanks her subjects, and her collaborators Dr. Charles Oman of MIT, Dr. Michael Zuschlag, and Andrew Kendra of the U.S. DOT Volpe Center, and also Prof. Thomas Sheridan, and Prof. James Kuchar, of MIT. She would also like to thank Carl Quesnel for his editorial assistance.

References

[1] A. H. Bellenkes, C. D. Wickens, and A. F. Kramer, "Visual scanning and pilot expertise: The role of attentional flexibility and mental model development," *Aviation, Space, and Environmental Medicine*, vol. 68, pp. 569-579, 1997.

[2] N. Moray, "Monitoring behavior and supervisory control," in *Handbook of Perception and Human Performance*, vol. II, K. R. Boff, L. Kaufman, and J. P. Thomas, Eds., New York, A Wiley-Interscience Publication, 1986.

[3] M. Hayashi, C. M. Oman, and M. Zuschlag, "Hidden Markov models as a tool to measure pilot attention switching during simulated ILS approaches," in *Proc. of 12th International Symposium on Aviation Psychology*, Dayton, OH, 2003, pp. 502-507.

[4] P. M. Fitts, R. E. Jones, and J. L. Milton, "Eye movements of aircraft pilots during instrument-landing approaches," *Aeronautical Engineering Review*, vol. 9, pp. 24-29, 1950.

[5] FAA, "Instrument Flying Handbook," US DOT Federal Aviation Administration, FAA-A-8083-15, 2001.

[6] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Englewood Cliffs, NJ: Prentice Hall, 1993, pp. 321-389.

[7] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford, UK: Clarendon Press, 1995, pp. 187-188.